1

NOTES ON THE THEORY OF
DYNAMIC PROGRAMMING—VII
TRANSPORTATION MODELS

By

Richard Bellman

P-771

28 November 1955

The RAND Corporation
1700 MAIN ST. · SANTA MONICA · CALIFORNIA

12

Summary

The purpose of this paper is to illustrate some applications of the functional equation technique of the theory of dynamic programming to a general class of problems arising in the study of networks, particularly those arising in transportation theory. ( )

NOTES ON THE THEORY OF DYNAMIC PROGRAMMING—VII
TRANSPORTATION MODELS

By
Richard Bellman

§1. Introduction

The problem of determining the structure of networks which are optimal with respect to different types of criteria is a burden shared in common by those who work in economic, industrial, organizational, communication, electronic, and computing fields. Relatively few of the variegated aspects of these difficult and important problems have been sampled to date, and none of these yield to simple analysis. For a brief survey of these questions see [1].

In this paper we wish to make a slight contribution to the general theory by showing the applicability of the functional equation technique of the theory of dynamic programming to some problems in transportation theory, and by analogy to the other fields mentioned above, as well.

The first problem we shall discuss is the Hitchcock-Koopmans transportation problem, cf. G. Dantzig, [4], and M. M. Flood, [5]. Next we shall treat a multi-stage version. Finally, we shall treat a closely related process arising in the studies of transportation systems by T. E. Harris and co-workers. Powerful computational algorithms exist for the solutions of these problems, some based directly or indirectly

upon the simplex method of Dantzig, and some upon characteristic features of the process, as, for example, the "flooding technique" of A. Boldyreff, [3], cf. also D. R. Fulkerson and L. R. Ford, [6].

The method we present is at the moment not applicable to all forms of the problems cited above, due to the limited capacity of present-day computing machines. Since these capacities are increasing with each year, the method will cover more and more cases as time goes on.

Even at the present time this method seems more efficient in treating certain special cases of the general processes, including some which occur in application. Furthermore, the method is equally applicable to processes where nonlinear and stochastic elements occur. In addition, the method is useful in determining the structure of optimal policies, the existence of "prices", and so on. These points will be discussed in a separate paper.

§2. The Hitchcock-Koopmans Transportation Problem

A problem which has been treated in considerable detail by a number of authors is the following:

"We are given a number of "sources", S_1, S_2, $\cdots$, S_i, $\cdots$, S_n, and a number of "sinks" or "terminals", T_1, T_2, $\cdots$, T_j, $\cdots$, T_N. Each source S_i has a quantity x_i of resources which must be transported to various of the terminals in such a way that the total quantity arriving at T_j fulfills a demand y_j. It is

assumed that $\sum_i x_i = \sum_j y_j$. Given the distances, d_{ij}, between the sources and the terminals, and assuming that the cost of shipping a unit quantity of resources between S_i and T_j is equal to d_{ij}, we wish to determine the routing which minimizes the total cost of supplying these demands".

One way of formulating this problem in analytic terms is the following: Let x_{ij} be the quantity of resources transported from S_i to T_j. Then

$$(1) \qquad \sum_j x_{ij} = x_i,$$

$$\sum_i x_{ij} = y_j, \qquad x_{ij} \geq 0,$$

and the total cost is given by

$$(2) \qquad C = \sum_{i,j} d_{ij} x_{ij}.$$

We wish to choose the quantities x_{ij}, subject to the constraints of (1), so as to minimize (2).

In this form, the problem has been solved numerically by means of various iterative techniques. Of these, the simplex method, with various modifications, seems most efficient.

§3. Dynamic Programming Formulation

We now wish to formulate the problem in dynamic programming terms. To do this, we regard the process as a multi-stage process in which we first fulfill the demands of T_1 from the

resources of the S_i, then the demands of T_2 from the remaining sources, and so on.

As we have mentioned above, one advantage of this formulation lies in the fact that we need no longer assume proportional costs, not valid in many situations because of the existence of certain "red-tape" or "set-up" costs, and we can, if we so desire, consider cases where the cost functions are stochastic.

For fixed demands, $y_1, y_2, \cdots, y_N$, it is clear that the minimum cost will be a function only of the quantities $x_1, x_2 \cdots, x_n$, and the number of terminals N. Let us then define the function

$$(1) \qquad f_N(x_1, x_2, \cdots x_n) = \operatorname*{Min}_{\{x_{ij}\}} C$$

To obtain a functional equation for $f_N(x)$, let us assume that we begin by allocating the quantities $x_{11}, x_{21}, \cdots, x_{n1}$ to supply the demand at T_1. Having done so, we have a problem of precisely the same type remaining with $N - 1$ terminals and quantities $x_1 - x_{11}, x_2 - x_{21}, \cdots, x_n - x_{n1}$ at the M sources. Hence, we obtain the functional equation

$$(2) \qquad f_N(x_1, x_2, \cdots, x_n) = \operatorname*{Min}_{\{x_{i1}\}} \left[\sum_{i=1}^{n} d_{i1}x_{i1} + f_{N-1}(x_1 - x_{11}, x_2 - x_{21}, \cdots, x_n - x_{n1}) \right],$$

where the minimum is over the region

$$(3) \qquad x_{i1} \geq 0, \quad \sum_{i=1}^{n} x_{i1} = y_1,$$

for $N \geq 2$, with

$$(4) \qquad f_1(x_1, x_2, \cdots, x_n) = d_{1N}x_1 + d_{2N}x_2 + \cdots + d_{nN}x_n.$$

The sequence f_N is now determined recurrently via (2).*

An important observation for computational purposes is that the dimension can always be reduced from n to n − 1 since

$$(5) \qquad \sum_{i=1}^{n} x_i = \sum_{j=1}^{N} y_j.$$

Hence, for fixed y_j, it is sufficient to specify the quantities $x_1, x_2, \cdots x_{n-1}$. Thus we may write

$$(6) \qquad f_N(x_1, x_2, \cdots, x_n) = f_N(x_1, x_2, \cdots, x_{n-1}).$$

Consequently, the case of two sources reduces to a one-dimensional problem, three sources to a sequence of two-dimensional maximization problems involving functions of two variables, and so on.

With the inevitable improvement in the "memory" of computing machines, we shall be able to handle the case of more and more sources in this fashion. The important point is that the utility of the method depends only upon the number of sources, while the number of terminals may be exceedingly large.

This method is particularly applicable when the functions f_N are desired for a large set of values of the x_i. This is the case if the x_i are stochastic variables and we wish to

---

*This is an application of the "principle of optimality", cf. [2]

determine the distribution of the minimum cost, or if we can determine the distribution of the x_i's so as to minimize the total cost, assuming that we have some degree of freedom in assigning the quantities at the S_i at the beginning of the process.

§4. A Multi-Stage Transportation Problem

Now consider the situation where we have a sequenc of sources

$$
(1)\qquad
\begin{array}{cccccc}
A_1 & A_2 & A_3 & & A_N & T_1 \\
\cdot & \cdot & \cdot & & \cdot & \cdot \\
B_1 & B_2 & B_3 & \cdots & B_N & T_2 \\
\cdot & \cdot & \cdot & & \cdot & \cdot \\
C_1 & C_2 & C_3 & & C_N & T_3 \\
\cdot & \cdot & \cdot & & \cdot & \cdot
\end{array}
$$

At the sources A_1, B_1, C_1 we have quantities x_1, y_1, and z_1 respectively, and at T_1, T_2, T_3 we have demands r_1, r_2, r_3 where

$$
(2)\qquad r_1 + r_2 + r_3 = \sum_i x_i + \sum_i y_i + \sum z_i .
$$

The process proceeds in the following manner. We write

$$
(3)\qquad
\begin{aligned}
x_1 &= x_{11} + x_{12} + x_{13} \\
y_1 &= y_{11} + y_{12} + y_{13} \\
z_1 &= z_{11} + z_{12} + z_{13}, \qquad x_{1j}, y_{1j}, z_{1j} \geq 0,
\end{aligned}
$$

where $x_{11} + y_{11} + z_{11}$ goes to A_2, $x_{12} + y_{12} + z_{12}$ to B_2 and $x_{13} + y_{13} + z_{13}$ to C_2. Starting with the new quantities $x_2 + x_{11} + y_{11} + z_{11}$ at A_2, $y_2 + x_{12} + y_{12} + z_{12}$ at B_2, $z_2 + x_{13} + y_{13} + z_{13}$ at C_2, the process continues in the same fashion.

Given the distances between the parts, the problem is to determine the routing which minimizes the total cost.

Let

(4) $f_k(x, y, z)$ = the cost incurred starting with x at A_k, y at B_k, and z at C_k, and employing an optimal policy.

Then $f_N(x, y, z)$ is determined as the solution of the usual Hitchcock-Koopmans transportation problem, and

$$(5) \qquad f_k(x, y, z) = \min_{\{x_{ij}\}} \Big[\sum_{1} (d_{11}^{(k)} x_{11} + d_{21}^{(k)} y_{11} + d_{31}^{(k)} z_{11}) + f_{k+1}(x + x_{11}^{(k)} + y_{11}^{(k)} + z_{11}^{(k)}, \ldots, z + x_{13}^{(k)} + y_{13}^{(k)} + z_{13}^{(k)})\Big],$$

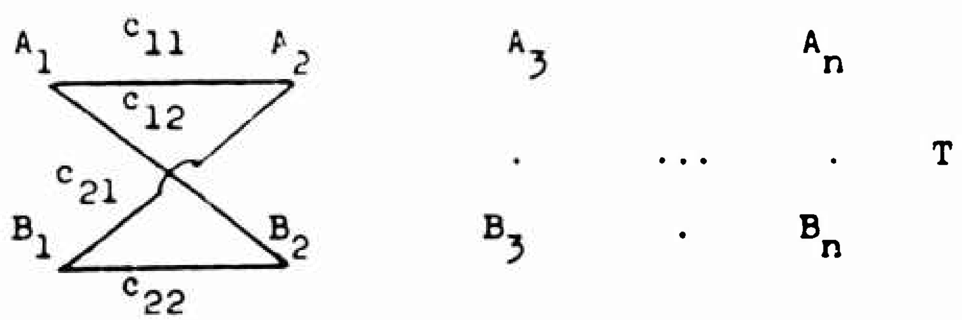
where the minimization is over the region described by (2).

§5. Railway and Communication Nets

An analogous class of problems arises from the study of railway and communication networks. Here we have networks

similar to those appearing in the previous section, with the difference that there are now capacity restraints on the flow between any two junctions and, sometimes, capacity restraints at the junctions.

Consider the following simple network:



Assume that trains at A_i and B_i can be sent either to A_{i+1} or B_{i+1} with the maximum flow between A_i and A_{i+1} given by $c_{11}^{(i)}$, between A_i and B_{i+1} by $c_{12}^{(i)}$, between B_i and A_{i+1} by $c_{21}^{(i)}$, and between B_i and B_{i+1} by $c_{22}^{(i)}$, for $i = 1, 2, \cdots, n - 1$, and finally the maximum flow from A_n to T given by d_1, and from B_n to T by d_2.

Starting with x trains at A_k and y trains at B_k, let

(1) $\quad f_k(x, y)$ = the number of trains arriving at T, using an optimal policy.

Clearly

(2) $$f_n(x, y) = \text{Min}\,(x, d_1) + \text{Min}\,(x, d_2),$$

and

(3) $$f_k(x, y) = \max_R \left[f_{k+1}(x_{11} + y_{21}, x_{12} + y_{22}) \right],$$

where the maximum is over the region

(4) (a) $x_{11} + x_{12} \leq x, \; y_{21} + y_{22} \leq y$

(b) $0 \leq x_{11} \leq c_{11}, \; 0 \leq x_{12} \leq c_{12},$

$0 \leq y_{21} \leq c_{21}, \; 0 \leq y_{22} \leq c_{22}.$

Bibliography

1. Bellman, R., "Mathematical Aspects of Scheduling Theory," RAND Paper P-651, 1955.

2. __________, "The Theory of Dynamic Programming," Bull. Amer. Math. Soc., Vol. 60 (1954), pp. 503-516.

3. Boldyreff, A., "Determination of the Maximal Steady State Flow of Traffic Through a Railroad Network", Jour. Operations Research Society, Vol. 3 (1955), pp. 443-466.

4. Dantzig, G., "Application of the Simplex Technique to a Transportation Problem", Activity Analysis of Production and Allocation", Wiley, 1953.

5. Flood, M. M., "On the Hitchcock Distribution Problem", Pac. Jour. Math., Vol. 3 (1953), pp. 369-396.

6. Fulkerson, D. R., and Ford, L., "Maximal Flow Through a Network", RAND Paper P-605, 1954.